

(12) INTERNATIONAL PUBLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property
Organization
International Bureau



Rec'd PCT/PTO

24 MAR 2005



(43) International Publication Date
8 April 2004 (08.04.2004)

PCT

(10) International Publication Number
WO 2004/029878 A1

(51) International Patent Classification⁷: **G06T 7/00, 5/50**

(21) International Application Number:
PCT/GB2003/004163

(22) International Filing Date:
25 September 2003 (25.09.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0222211.5 25 September 2002 (25.09.2002) GB

(71) Applicant (for all designated States except US):
FORTKEY LIMITED [GB/GB]; Elvingston Science
Centre, Elvingston, Gladsmuir, East Lothian EH33 1EH
(GB).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **MORRISON,
Steven** [GB/GB]; C/o Fortkey Limited, Elvingston Sci-
ence Centre, Elvingston, Gladsmuir, East Lothian EH33
1EH (GB). **CLARKE, Stuart, James** [GB/GB]; C/o
Fortkey Limited, Elvingston Science Centre, Elvingston,

Gladsmuir, East Lothian EH33 1EH (GB). **LINNETT,
Laurence, Michael** [GB/GB]; C/o Fortkey Limited,
Elvingston Science Centre, Elvingston, Gladsmuir, East
Lothian EH33 1EH (GB).

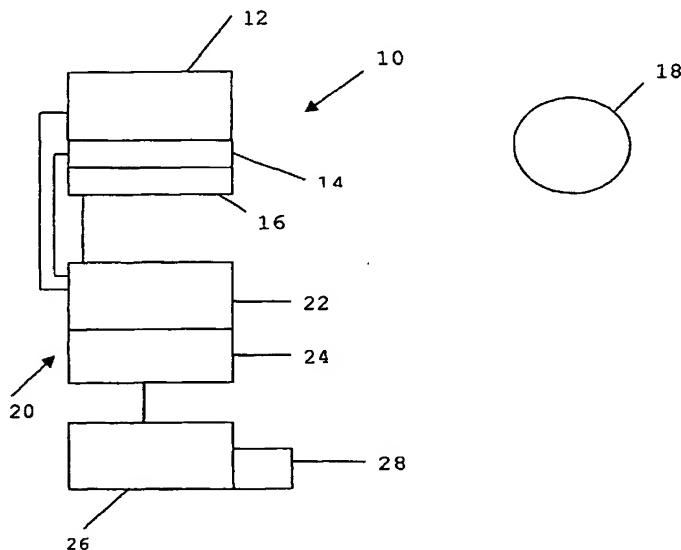
(74) Agent: **KENNEDYS PATENT AGENCY LIMITED**;
Floor 5, Queens House, 29 St. Vincent Place, Glasgow G1
2DT (GB).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC,
SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA,
UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE,
ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO,
SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM,
GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: IMAGING AND MEASUREMENT SYSTEM



(57) Abstract: Apparatus and method for presenting a highly spatially accurate visualisation of a scene from which measurements can be taken. A sensor is located in relation to a camera, and provides positional characteristics of the camera as it collects frames of video images. Using the positional characteristics the frames are corrected. The corrected frames are then synchronised to form an accurate mosaic of a scene. Example embodiments are described where the moving camera is used to survey or inspect underwater apparatus, roads, runways, railways, crime or accident scenes, archaeological digs and the inside of boilers, chimneys and pipelines.



Published:

- with international search report
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

1 IMAGING AND MEASUREMENT SYSTEM

2

3 The present invention relates to video mosaicing and, in
4 particular, to a method and system for providing a highly
5 spatially accurate visualisation of a scene from which
6 measurements can be taken.

7

8 A video mosaic is a composite image produced by stitching
9 together frames from a video sequence such that similar
10 regions overlap. The output gives a representation of
11 the scene as a whole, rather than a sequential view of
12 parts of that scene, as in the case of a video survey of
13 an area. One of the best known applications of this
14 technique being the creation of panoramic photographs of
15 a scene.

16

17 In publishing and image retouching applications the
18 mosaics are manually generated which is a costly and time
19 consuming process. More recently a system for
20 automatically generating a mosaic has been suggested, US
21 Patent 5,649,032, which provides the possibility for
22 real-time video mosaicing. This Patent details
23 applications for display of an image, compression of an
24 image for storage and when constructed, to a surveillance

1 system suitable for determining enemy movement on a
2 battlefield, a burglar entering a warehouse, and the
3 like.

4

5 Video mosaics constructed in this fashion are not suited
6 to applications involving the making of accurate
7 measurements for the following reasons.

8

9 Firstly, it is vital to perform a camera calibration
10 procedure to estimate and hence correct for the
11 distortions caused by the internal geometry of the
12 camera. Uncorrected, these distortions will significantly
13 degrade the accuracy of any measurements made from the
14 mosaic.

15

16 Secondly, the nature of the accumulation of errors in the
17 estimation of rotations between frames leads a drift
18 characteristic of a "random walk" which will seriously
19 degrade the accuracy of long range measurements.

20

21 Finally, non-translational changes in the camera position
22 (e.g. pitch and roll) will lead to perspective changes
23 between frames which will also degrade the positional
24 accuracy of the constructed mosaic. Although it is
25 possible to estimate the variation in camera attitude
26 from the video frames, the accumulation of the associated
27 errors would again lead to degradation in measurement
28 accuracy.

29

30 It is an object of the present invention to provide a
31 measurement system and method using video mosaicing which
32 obviates or mitigates at least some of the disadvantages
33 in the prior art.

34

1 It is further object of at least one embodiment of the
2 present invention to provide a measurement system and
3 method to provide a highly spatially accurate
4 visualisation of a scene from which measurements can be
5 taken.

6
7 It is a still further object of at least one embodiment
8 of the present invention to provide a measurement system
9 and method from which one can make measurements of a
10 scene to millimetre resolution.

11
12 According to a first aspect of the present invention
13 there is provided apparatus for presenting a highly
14 spatially accurate visualisation of a scene from which
15 measurements can be taken, the apparatus comprising:

16
17 at least one camera for recording a plurality of
18 frames of video images of the scene;

19
20 at least one sensor mounted in relation to the
21 camera for recording sensor data on positional
22 characteristics of the camera as the at least one
23 camera is moved with respect to the scene; and

24
25 image processing means including a first module for
26 synchronising the frames with the sensor data to
27 form corrected frames; and a second module for
28 constructing an accurate mosaic from the corrected
29 frames.

30
31 By first correcting the video frames prior to the
32 mosaiced image being formed, distortions present in the
33 frames recorded by the one or more cameras can be removed

1 and so enhance the spatial resolution over the entire
2 mosaiced image.

3

4 Preferably the at least one camera is a video camera
5 capturing 2 dimensional digital images.

6

7 The at least one sensor may comprise any sensor capable
8 of making a positional measurement. Preferably the at
9 least one sensor comprise sensors making a measurement
10 relating to attitude or distance. Preferably also the at
11 least one sensor comprises a digital compass.
12 Advantageously the digital compass records roll, pitch
13 and yaw. Preferably also, the at least one sensor
14 comprises an altimeter and/or bathymetric sensor.

15

16 Advantageously the camera(s) and sensor(s) are mounted on
17 a moving platform. In use the platform may be mounted on
18 a vehicle to allow movement of the camera(s) and
19 sensor(s) over or through the scene to be imaged.

20

21 The apparatus may further include a calibration system
22 from which the at least one camera is calibrated. In this
23 way spherical lens distortion e.g. pincushion distortion
24 and barrel distortion can be corrected prior to use of
25 the camera(s). Further non-equal scaling of the pixels in
26 the x and y axis is corrected together with a skew of the
27 two image axis from the perpendicular.

28

29 Advantageously the calibration system includes a
30 chessboard pattern or regular grid. This provides for
31 multiple images to be taken from multiple viewpoints so
32 that the distortions can be estimated and compensated
33 for.

34

1 Preferably the first module performs a perspective
2 correction to the images using the sensor data.
3 Preferably also, the corrected frames are of a
4 preselected position with reference to the scene.
5 Optionally the corrected frames may be of preselected
6 attitude and distance.

7

8 Preferably the second module accomplishes video mosaicing
9 via a correlation technique based on frequency contents
10 of the images being compared.

11

12 Preferably the apparatus further includes display means
13 for providing a visual image of the mosaic. Preferably
14 also the apparatus further comprises data storage means
15 to allow the mosaic to be stored for viewing at a later
16 time.

17

18 Preferably also the apparatus includes a graphic user
19 interface (GUI). More preferably the GUI is included with
20 the display system. Advantageously the GUI includes means
21 to allow a user to select and make measurements between
22 points in the visual image of the mosaic. Optionally the
23 GUI provides a user with means to control the movement of
24 the at least one camera.

25

26 According to a second aspect of the present invention
27 there is provided a method for presenting a highly
28 spatially accurate visualisation of a scene from which
29 measurements can be taken, the method comprising the
30 steps;

31

32 (a) recording a plurality of frames of video images
33 of the scene from a camera;

1 (b) recording sensor data on positional
2 characteristics of the camera as the camera is
3 moved with respect to the scene;

4 (c) synchronising the frames with the sensor data
5 to form corrected frames; and

6 (d) constructing an accurate mosaic from the
7 corrected frames.

8

9 Preferably the method includes the step of calibrating
10 the camera prior to step (a). This calibration may remove
11 distortion effects within the camera.

12

13 Preferably the step of calibrating includes the step of
14 taking multiple images of a chessboard pattern or regular
15 grid from multiple viewpoints and further estimating and
16 compensating for the distortions.

17

18 Preferably the synchronisation step includes the step of
19 performing a perspective correction to the images using
20 the sensor data.

21

22 Preferably also the step of video mosaicing is achieved
23 using a correlation technique based on frequency contents
24 of the images being compared.

25

26 Preferably the method further includes the step of
27 providing a visual image of the mosaic.

28

29 Advantageously the method further includes the step of
30 taking a measurement from the visual image.

31

32 Optionally the method may include the step of storing the
33 images so that they may be accessed by spatial position.

34

1 This method may advantageously be used to record crime
2 scenes, accident scenes, archaeological digs and the like
3 where traditional methods of image recordal and distance
4 measurement are time consuming. Additionally by storing
5 the mosaiced images, distances previously not measured
6 within the scene can be regenerated and accurately
7 measured without having to reconstruct or preserve the
8 original scene.

9
10 According to a third aspect of the present invention
11 there is provided a method of performing a survey in a
12 fluid, the method comprising the steps of;

- 13
- 14 (a) mounting a camera and a plurality of sensors on a
15 platform capable of movement in the fluid;
 - 16 (b) moving the platform through the fluid while
17 recording visual images on the camera and taking
18 sensor data relating to the attitude and distance
19 of the platform from objects of interest within
20 the fluid;
 - 21 (c) synchronising the visual images to the sensor data
22 to provide corrected visual images relating to a
23 fixed distance and attitude;
 - 24 (d) video mosaicing the images to form an accurate
25 video mosaic as a visual image of the scene
26 surveyed.

27
28 Preferably the method includes the step of precalibrating
29 the camera to compensate for distorting artefacts
30 inherent within the camera.

31
32 Preferably the method includes the step of displaying the
33 visual image. More preferably the method includes the
34 step of taking a measurement from the visual image.

1

2 Preferably the fluid is water, so that measurements can
3 be made underwater. In this way pipe spool dimensions
4 can be taken underwater as can determination be made of
5 the degree of damage or degradation of pipelines.

6

7 Advantageously the platform may be mounted on an
8 autonomous underwater vehicle (AUV) or a remotely
9 operated vehicle (ROV). Alternatively the platform may
10 be mounted on a PIG (pipeline inspection gauge), so that
11 the camera can be moved through a pipeline to inspect the
12 inner surface of the pipeline.

13

14 Preferably the method includes the step of storing the
15 mosaiced images for viewing later.

16

17 Embodiments of the present invention will now be
18 described, by way of example only, with reference to the
19 following Figures, of which:

20

21 Figure 1 is a schematic diagram of a first
22 embodiment of the present invention;

23

24 Figure 2 is a schematic diagram of a second
25 embodiment of the present invention;

26

27 Figure 3 is a flow diagram depicting the stages of
28 the sensor data integration with the algorithms
29 required for the construction of the measurement
30 mosaic of the second embodiment;

31

32 Figure 4 depicts a schematic of the camera pose
33 alteration required to correct for perspective in

each of the image frames by application of the pitch and roll sensor data in the second embodiment;

Figure 5 shows a flow diagram of the method applied when correcting images for the sensor roll and pitch data concurrently with the camera calibration correction as in the second embodiment;

Figure 6 is a schematic diagram of a third embodiment of the present invention; and

Figure 7 is a schematic diagram of a fourth embodiment of the present invention.

Referring initially to Figure 1 there is shown imaging apparatus, generally indicated by reference numeral 10, according to a first embodiment of the present invention. Apparatus 10 comprises a camera 12 mounted with sensors 14,16. The camera 12 captures a series of frames of video images as the camera 12 and sensors 14,16 are moved over an object 18. During this movement the sensors 14,16 record data on the attitude and distance of the camera 12 from the object 18. The sensor data and video images are input an image processor, generally indicated at 20. The processor 20 includes a first module 22 in which the frames are synchronised with the sensor data, as will be described hereinafter. The first module 22 outputs corrected video image from which is constructed a video mosaic in the second module 24, as described hereinafter. The video mosaic of the object 18 is displayed on a monitor 26 of a personal computer. Using a graphical user interface 28 of the personal computer a user can select points on the video mosaic and obtain distance measurements of the object 18. The measurements provide

1 millimetre accuracy over 20 metre distances to the
2 object. This is achieved by correcting variations in
3 pixel dimensions with the sensor data and/or camera
4 calibration, described hereinafter, and using the sensor
5 data to also provide a determination of pixel dimensions
6 in terms of real metric units.

7
8 Figure 2 depicts a schematic diagram of a second
9 embodiment of the present invention illustrating the
10 hardware and the high level processes. This embodiment
11 consists of an instrumented camera platform, generally
12 indicated by reference numeral 30, incorporating a video
13 camera 32 which may be analogue or digital, a digital
14 compass 34 and an altimeter sensor 36. The sensors 34,36
15 measure the attitude (roll, pitch and yaw/heading) of the
16 platform 30 and the distance from the camera platform 30
17 to an object being viewed. In underwater applications,
18 an additional bathymetric sensor may be used to measure
19 the depth of submergence of the camera platform 30. Thus
20 the platform 30 will be mounted on a suitable vehicle 35
21 e.g. underwater remotely operated vehicle (ROV), aircraft
22 or even a hand-held mounting and moved across the scene
23 of interest. As in the first embodiment, the video and
24 sensor data is made available to the operator 37 of the
25 system for live display. Additionally, the video and
26 sensor data is stored 38 in a format which allows precise
27 synchronization between the video and sensor data. The
28 stored data 38 may be retrieved and used to construct a
29 video mosaic image 40 representing a plan view of the
30 scene being surveyed where pixel scale is maintained
31 throughout the image. During the construction of this
32 mosaic image corrections are applied to the video frames
33 to correct the inherent distortions due to the video
34 camera and to compensate for the effects of camera

1 platform attitude and distance to the viewed scene.
2 These corrections ensure that the constructed mosaic
3 image 40 is an accurate representation of the scene being
4 surveyed, with the relative scales and positions of the
5 objects contained within the scene being preserved as
6 well as possible. Once constructed, it is possible to
7 obtain measurements 42 of objects contained within the
8 mosaic image using a graphical user interface.

9
10 Figure 3 depicts a flow diagram of the stages required to
11 construct the video mosaic image. The first stage in
12 this process is to acquire a frame of video data 50 and
13 the corresponding sensor data 52 for this frame, from the
14 storage unit 38. The video frame 50 is then corrected to
15 compensate for the effects of the camera distortion and
16 the camera platform attitude 54. This stage requires
17 knowledge of the camera internal parameters which are
18 estimated by a calibration method described later, and
19 the pitch and roll angles 56 recorded by the digital
20 compass 34. The corrected image 58 is then input into
21 the mosaicing procedure 60 where it is compared with the
22 previous corrected video frame 50 in the video sequence.
23 This procedure attempts to estimate the translation in x
24 and y axes between the two frames by comparing the
25 correlations between the frames in the frequency domain.
26 The rotation between frames and the scale change between
27 frames is determined from the compass heading and
28 altitude/depth information 62. The next stage 64 is to
29 apply the transformation parameters to the new frame and
30 incorporate it into the final mosaic image 66, a process
31 known as "stitching". Finally the pixel size may be
32 determined by the use of a calibration target placed in
33 the scene, or directly from the camera calibration
34 parameters and altimeter sensor data.

1

2 We shall consider the steps taken in the method in more
3 detail. Beginning with the camera 32, all cameras suffer
4 from various forms of distortion. This distortion arises
5 from certain artefacts inherent to the internal camera
6 geometric and optical characteristics (otherwise known as
7 the intrinsic parameters). These artefacts include:

8

9 (a) spherical lens distortion about the principal
10 point of the system. The two common definitions
11 for this type of distortion are pincushion
12 distortion and barrel distortion;

13

14 (b) non-equal scaling of pixels in the x and y-axis.
15 This is arrived at through the estimation of the
16 effective camera focal length in both the x and y
17 pixel scales; and

18

19 (c) a skew of the two image axes from the
20 perpendicular.

21

22 For high accuracy mosaicing the parameters leading to
23 these distortions must be estimated and compensated for.
24 In order to correctly estimate these parameters images
25 taken from multiple viewpoints of a regular grid, or
26 chessboard type pattern are used. The corner positions
27 are located in each image using a corner detection
28 algorithm. The resulting points are then used as input
29 to a camera calibration algorithm as well documented in
30 the literature.

31

32 The estimated intrinsic parameter matrix A is of the form

$$A = \begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

1

2

3 where α and β are the focal lengths in x and y pixels
4 respectively, γ is a factor accounting for skew due to
5 non-rectangular pixels, and (u_0, v_0) is the principle point
6 (that is the perpendicular projection of the camera focal
7 point onto the image plane).

8

9 During the creation of the mosaic, the integration of the
10 sensor data is performed in two phases; as is illustrated
11 in Figure 4. The first of these involves the use of the
12 pitch and roll measurements 56 from the compass 34 to
13 perform a perspective correction on each of the frames
14 prior the mosaicing procedure 60. A diagram showing the
15 situation modelled by this correction is provided in
16 figure 4. When correcting for perspective the new camera
17 position 70 is at the same height 72 as the original
18 viewpoint 74, not the slant range distance 76a,b,c. Thus
19 any correction for perturbations in pitch or roll will
20 not be misinterpreted as a change in camera height, which
21 may be considered either as a separate process handled
22 within the mosaicing procedure 60 itself, or gained from
23 the bathymetric sensor readings.

24

25 This perspective correction 54 is performed concurrently
26 with the camera calibration correction 55 following the
27 steps outlined in Figure 5. Figure 5 illustrates the
28 steps applied to all pixel positions in the corrected
29 image 58. Starting with the corrected image pixel
30 position 58, we obtain the corresponding pixel position
31 in the cameras true reference frame 82, we then obtain

1 the position in captured image distorted by the camera
2 calibration parameters 84, interpolate for value at
3 resulting subpixel level 86 and insert interpolate value
4 into initial corrected image pixel position 88.

5

6 Concatenating these two operations in this way saves on
7 both processing time and memory requirements. These
8 processes combine mathematically in the following way:

9

10 If \underline{u} is the corrected pixel position, the corresponding
11 position in the reference frame of the camera, normalised
12 according the camera focal length in y pixels (β) and
13 centred on the principle point (u_0, v_0) , is
14 $\underline{c}' = [(c_1'', c_2'', c_3'') / c_4'' - (u_0, v_0)] / \beta$ where $\underline{c}'' = PR_y R_x P^{-1} \underline{u}$. The pitch
15 and roll are represented by the rotation matrices R_x and
16 R_y respectively, with P being the perspective projection
17 matrix which maps real world coordinates onto image
18 coordinates. Following this the pixel position in the
19 captured image is calculated as $\underline{c} = A \tau_c \underline{c}'$. The scalar τ_c
20 represents the radial distortion applied at the camera
21 reference frame coordinate \underline{c}' . The matrix A is as
22 defined previously.

23

24 In estimating interframe mosaicing parameters of video
25 sequences there are currently two types of method
26 available. The first uses feature matching within the
27 image to locate objects and then to align the two frames
28 based on the positions of common objects. The second
29 method is frequency based, and uses the properties of the
30 Fourier transform.

31

1 Given the volume of data involved (a typical capture rate
2 being 25 frames per second) it is important that we
3 utilise a technique which will provide a fast data
4 throughput, whilst also being highly accurate in a
5 multitude of working environments. In order to achieve
6 these goals, the preferred embodiment employs the
7 correlation technique based on the frequency content of
8 the images being compared. This approach has two main
9 advantages; firstly, regions which would appear
10 relatively featureless, that is those not containing
11 strong corners, linear features, and such like, still
12 contain a wealth of frequency information representative
13 of the scene. This is extremely important when mosaicing
14 regions of the seabed for example, as definite features
15 (such as corners or edges) may be sparsely distributed;
16 if indeed they exist at all; and secondly, the fact that
17 this technique is based on the Fourier transform means
18 that it opens itself immediately to fast implementation
19 through highly optimized software and hardware solutions.

20

21 The second phase of integration is applied in tandem with
22 the frequency correlation technique and incorporates both
23 the altimeter and heading readings.

24

25 The mosaicing technique is capable of estimating the
26 rotations between adjacent frames in the mosaic to an
27 extremely high degree of accuracy. Unfortunately, the
28 nature of the accumulation of the errors corresponds to a
29 stochastic process called a "random walk". This has the
30 effect of leading to a drift in the estimated track. For
31 short range mosaics this effect is limited and may be
32 discounted, thus allowing use of Fourier rotation
33 measurements. However, for long range mosaics this will
34 not be the case. In order to overcome this, the yaw data

1 is utilised from the digital compass to provide a stable
2 reference for the camera heading. This greatly increases
3 the overall accuracy of the reconstructed mosaic.

4
5 For each image comparison, the interframe rotation and
6 scaling values are obtained from the difference in the
7 heading and bathymetric readings for that image pair.
8 The second image is then corrected to the same
9 orientation and scale of the first. This way only the
10 translation in x and y pixels need be estimated. Having
11 obtained the necessary parameters of the differences in
12 position of the two images, they can be placed in their
13 correct relative positions. The next frame is then
14 analysed in a similar manner and added to the evolving
15 mosaic image.

16
17 We shall now give a description of the implementation
18 procedures used in this invention for translation
19 estimation in Fourier space.

20
21 In Fourier space, translation is a phase shift. We
22 therefore must utilise the differences in the phase to
23 determine the translational shift. Let the two images be
24 described by $f_1(x,y)$ and $f_2(x,y)$ where (x,y) represents a
25 pixel at this position. Then for a translation (dx,dy) the
26 two frames are related by

27

$$28 \quad f_2(x,y) = f_1(x+dx, y+dy)$$

29

30 The Fourier transform magnitudes of these two images are
31 the same since the translation only affects the phases.
32 Let our original images be of size $(cols,rows)$, then each of
33 these axes represents a range of 2π radians. So a shift

1 of dx pixels corresponds to $2\pi dx/cols$ shift in phase for
2 the column axis. Similarly, a shift of dy pixels
3 corresponds to $2\pi dy/rows$ shift in phase for the row axis.

4

5 To determine a translation, we Fourier transform the
6 original images, compute the magnitude (M) and phases
7 (ϕ) of each of the pixels and subtract the phases of each
8 pixel to get $d\phi$. We then take the average of the
9 magnitudes (they should be the same) and the phase
10 differences and compute a new set of real (\Re) and
11 imaginary (\Im) values as $\Re = M \cos(d\phi)$ and $\Im = M \sin(d\phi)$. These
12 (\Re, \Im) values are then inverse Fourier transformed to
13 produce an image. Ideally, this image will have a single
14 bright pixel at a position (x, y) , which represents the
15 translation between the original two images, whereupon a
16 subpixel translation estimation may be made.

17

18 It is not always that case that the peak is unique
19 however. When we have translation close to zero, the
20 gained true peak is often distorted by a secondary peak
21 at the origin. For this reason we place a lower
22 acceptance bound on the translation. If the gained
23 translation is lower than this, then the current new
24 frame is discarded, and the next is compared to the same
25 initial frame. This process has the added speed
26 advantage that frames are only stitched into the mosaic
27 if a reasonable translation has occurred.

28

29 A final point to note concerning this technique is that
30 we must first window the intensity values to be Fourier
31 transformed, ensuring that they are reduced to zero at
32 the boundary. This removes the step discontinuities at
33 the boundaries, making the periodic image, implied when

1 stepping into the Fourier domain, appear continuous in
2 all directions.

3

4 Following acquisition of the interframe mosaicing
5 parameters it remains for the video images to be stitched
6 into a single mosaic so that measurements between imaged
7 positions may be achieved. This is performed using a
8 similar philosophy to that adopted when correcting for
9 perspective and camera calibration. Given a pixel
10 position within the mosaic, what was the corresponding
11 sub-pixel position in the original frame? The
12 construction of the mosaic is also performed in such a
13 way as to minimise the amount of memory required to
14 contain the result.

15

16 In order to determine this mapping we first generate the
17 camera track file containing the frame centre positions,
18 orientations, and scale factors from the parameter file
19 output by the mosaicing algorithm. This is done through
20 accumulation of local translations, rotations, and
21 scaling factors, each having undergone a rotation and
22 scaling to make them local to the mosaic reference frame.

23

24 Following this, we may calculate the coordinates of the
25 i^{th} frame pixel position (x_{f_i}, y_{f_i}) , in terms of the
26 corresponding mosaic pixel position (x_m, y_m) , as

27

$$28 \quad \begin{bmatrix} x_{f_i} \\ y_{f_i} \end{bmatrix} = \frac{1}{z_i} \begin{bmatrix} \cos(\theta_i) & -\sin(\theta_i) \\ \sin(\theta_i) & \cos(\theta_i) \end{bmatrix} \begin{bmatrix} x_m - \frac{\rho_{c_i} - 1}{2} \\ y_m - \frac{\rho_{r_i} - 1}{2} \end{bmatrix} + \begin{bmatrix} \frac{f_c - 1}{2} \\ \frac{f_r - 1}{2} \end{bmatrix}$$

29

30 where θ_i and z_i are the rotation and scaling values which
31 place the i^{th} frame into the mosaic, the size of area

1 required to fully contain the frame in the mosaic is
2 $\rho_c \times \rho_r$ pixels, and the original frame size is $f_c \times f_r$
3 pixels. We then interpolate the sub-pixel value at
4 position (x_{fi}, y_{fi}) in frame i , and place this value into
5 mosaic pixel position (x_m, y_m) .

6
7 Given the stitched mosaic it remains to make a
8 measurement between selected points in the final result.
9 In order to accomplish this, the pixel size must be
10 determined through use of either a calibration target
11 placed in the scene, or through use of the camera
12 calibration parameters and altimeter sensor data.
13 Following this calibration, the distance in pixels
14 between the selected points is multiplied by the true
15 distance subtended by each pixel to provide an accurate
16 length measurement.

17
18 The apparatus and method of the present invention lends
19 itself to the following applications particularly as
20 applied to underwater surveying:

21
22 (a) Metrology, through the measurement of physical
23 dimensions in difficult to access environments;

24
25 (b) Geo-referencing - in conventional video surveys
26 the data is stored in a video format where each
27 part of the survey is accessed by frame number.
28 Under the present invention a survey can be
29 stored as one or more mosaiced images which can
30 advantageously be accessed by spatial position
31 and integrated with other geo-referenced data
32 such as maps, sidescan sonar, and engineering
33 drawings;

1
2 (c) Video compression - while video recording of a
3 survey requires vast storage capacity and leads
4 to data being stored on difficult to access
5 magnetic tape media or in compressed forms on a
6 computer, the present invention provides a
7 compact data size as redundant information when
8 images overlap is removed. This is done with very
9 little degradation to the image quality compared
10 to video compression methods. It is also possible
11 to reconstruct a video of the original video
12 survey; and

13
14 (d) Navigation as the video mosaicing process
15 involves the measurement of translations
16 rotations and scalings that are present in the
17 video sequence, the apparatus can provide
18 navigational information about the platform on
19 which it may be mounted. As the navigational
20 information extracted from the video sequence may
21 be extremely accurate (<1cm) over short ranges,
22 the information can be used to aid positioning of
23 equipment, station holding and offers a potential
24 benefit to the development of a synthetic
25 aperture sonar system.

26
27 It will be appreciated that the second embodiment could
28 be adapted to inspect ships' hulls in order to check for
29 hull integrity or the prevention of smuggling or
30 terrorist threats. In this application the camera(s) and
31 sensors are mounted onto a remotely operated vehicle
32 (ROV) which is used to scan the hull of the ship. In
33 this configuration, the sensors could include an
34 altimeter to measure distance between the camera and ship

1 hull, and a digital compass unit to measure the platform
2 attitude. The sensor data can be used to apply scaling
3 and perspective corrections respectively to the camera
4 frames, prior to mosaicing the video frames into a large
5 image. The mosaic image may be used to identify the
6 position of any area of interest on the ship's hull.

7

8 A further application of this methodology is that of
9 internal pipe-like structure inspection, where pipe-like
10 structures include pipelines, boilers, and chimneys for
11 example. In this embodiment a system 100 includes a
12 plurality of cameras 90 are placed in a circular
13 arrangement as shown in figure 6 to provide a 360 degree
14 field of view, and images gathered of the surrounding
15 surface 92. Lighting sources 94 are placed adjacent to
16 the cameras 90; suitably illuminating the surface 92
17 being inspected. The cameras 90 are synchronised with
18 images gathered instantaneously being distortion
19 corrected depending on the camera calibration parameters,
20 arrangement of the cameras, and position of the camera
21 system within the pipe structure, thereby providing
22 images from which the accurate measurements of distances
23 along the pipe sidewall 92 may be obtained. The position
24 within the structure can be determined by separate range
25 finding sensors 96 mounted locally to each camera and
26 synchronised with that camera, these supply the distance
27 to the pipe structure sidewall of that camera. Via a
28 processor 98 the instantaneously grabbed images are then
29 accumulated into a mosaiced image strip containing the
30 entire imaged surface at that particular moment in time.
31 The system 100 can be propelled through a boiler or pipe
32 like structure via any means including gravity (a
33 vertical pipeline or chimney for example), a pulley
34 system pulling/pushing the setup, or by attaching to the

1 camera rig an arrangement of support struts with wheels,
2 these may be motorised or pushed/pulled through the pipe
3 structure by some external means. As the number of
4 strips accumulates over time they are automatically
5 stitched to form a mosaic of the surface under
6 inspection; the inside of a pipe, chimney, or boiler.

7
8 A yet further application of an embodiment of invention
9 described here is in the inspection of roads, runways and
10 railway lines. In this embodiment the system 102 could
11 consist of video cameras 104 mounted on a suitable
12 vehicle 106 facing towards the ground with the addition
13 of suitable lighting 108 to illuminate the surface being
14 inspected. In this configuration the additional sensors
15 could include a GPS receiver 110 that can be used to
16 provide additional global positioning information
17 synchronised to the video data. The video frames will be
18 corrected for camera and perspective distortion prior to
19 input to the mosaicing operation in the processor 112. A
20 video mosaic constructed from the combined (in the case
21 of more than one camera) and corrected video frames will
22 be generated. This image may be used to identify and
23 measure surface defects and to determine global positions
24 of these defects. The incorporation of GPS positional
25 information can further enable the generated mosaic image
26 to be referenced to a geographical information system
27 (GIS).

28
29 The main advantage of the present invention is that it
30 provides a video mosaic image from which measurements
31 with millimetre accuracy can be taken. High spatial
32 resolution is attainable by fusing the sensor data with
33 the video images and then reconstructing the mosaic from
34 a selected reference point. This allows measurements to

1 be made from the video mosaic as the pixel dimensions are
2 provided in terms of metric units scaled from the objects
3 being surveyed. Use of a correlation technique based on
4 the frequency content of the images being compared
5 provides the advantages of allowing imaging of generally
6 featureless scenes such as the seabed and as the
7 technique is based on the Fourier Transform the data can
8 be processed in real time through the implementation of
9 highly optimised software and hardware solutions.

10

11 Further the present invention provides advantages over
12 traditional ways of obtaining measurements. Firstly, it
13 may be used in environments where it is either hazardous
14 or difficult to use conventional manual measurement
15 methods. For example the measurement of pipeline spool
16 pieces on the seafloor, can be conducted by mounting the
17 camera and sensors on an ROV which can be flown over the
18 two ends of the pipeline to be connected by the spool
19 piece. Currently a method involving triangulation of
20 acoustic transceivers is employed for this application.
21 This is a time consuming method which requires the use of
22 divers and some expert knowledge. A second advantage is
23 that in the case of scenes containing a number of objects
24 that must have their positions or separations recorded, a
25 survey can be conducted and the measurements made at a
26 later time, with the minimum of delay incurred at the
27 scene. This would be a considerable benefit in recording
28 accident scenes or archaeological digs.

29

30 It will be appreciated by those skilled in the art that
31 various modifications may be made to the invention herein
32 described without departing from the scope thereof.

1 CLAIMS

2

3 1. Apparatus for presenting a highly spatially accurate
4 visualisation of a scene from which measurements can
5 be taken, the apparatus comprising:

6

7 at least one camera for recording a plurality of
8 frames of video images of the scene;

9

10 at least one sensor mounted in relation to the
11 camera for recording sensor data on positional
12 characteristics of the camera as the at least one
13 camera is moved with respect to the scene; and

14

15 image processing means including a first module for
16 synchronising the frames with the sensor data to
17 form corrected frames; and a second module for
18 constructing an accurate mosaic from the corrected
19 frames.

20

21 2. Apparatus as claimed in Claim 1 wherein the at least
22 one camera is a video camera capturing 2 dimensional
23 digital images.

24

25 3. Apparatus as claimed in Claim 1 or Claim 2 wherein
26 the at least one sensor comprises a sensor capable
27 of making a positional measurement.

28

29 4. Apparatus as claimed in Claim 3 wherein the at least
30 one sensor comprises a digital compass.

31

32 5. Apparatus as claimed in Claim 3 or Claim 4 wherein
33 the at least one sensor comprises an altimeter
34 and/or bathymetric sensor.

1

2 6. Apparatus as claimed in any preceding Claim wherein
3 the camera(s) and sensor(s) are mounted on a moving
4 platform.

5

6 7. Apparatus as claimed in any preceding Claim wherein
7 the apparatus further includes a calibration system
8 from which the at least one camera is calibrated.

9

10 8. Apparatus as claimed in any preceding Claim wherein
11 the first module performs a perspective correction
12 to the images using the sensor data.

13

14 9. Apparatus as claimed in any preceding Claim wherein
15 the second module accomplishes video mosaicing via a
16 correlation technique based on frequency contents of
17 the images being compared.

18

19 10. Apparatus as claimed in any preceding Claim wherein
20 the apparatus further includes display means for
21 providing a visual image of the mosaic.

22

23 11. Apparatus as claimed in any preceding Claim wherein
24 the apparatus further comprises data storage means
25 to allow the mosaic to be stored.

26

27 12. Apparatus as claimed in any preceding Claim wherein
28 the apparatus includes a graphic user interface
29 (GUI).

30

31 13. A method for presenting a highly spatially accurate
32 visualisation of a scene from which measurements can
33 be taken, the method comprising the steps;

34

- 1 (a) recording a plurality of frames of video images
2 of the scene from a camera;
- 3 (b) recording sensor data on positional
4 characteristics of the camera as the camera is
5 moved with respect to the scene;
- 6 (c) synchronising the frames with the sensor data
7 to form corrected frames; and
- 8 (d) constructing an accurate mosaic from the
9 corrected frames.
- 10
- 11 14. A method as claimed in Claim 13 wherein the method
12 includes the step of calibrating the camera prior to
13 step (a).
- 14
- 15 15. A method as claimed in Claim 13 or Claim 14 wherein
16 the synchronisation step includes the step of
17 performing a perspective correction to the images
18 using the sensor data.
- 19
- 20 16. A method as claimed in any one of Claims 13 to 15
21 wherein the step of video mosaicing is achieved
22 using a correlation technique based on frequency
23 contents of the images being compared.
- 24
- 25 17. A method as claimed in any one of Claims 13 to 16
26 wherein the method further includes the step of
27 providing a visual image of the mosaic.
- 28
- 29 18. A method as claimed in any one of Claims 13 to 17
30 wherein the method further includes the step of
31 taking a measurement from the visual image.
- 32
- 33 19. A method as claimed in any one of Claims 13 to 18
34 wherein the method includes the step of storing the

1 images so that they may be accessed by spatial
2 position.

3

4 20. A method of performing a survey in a fluid, the
5 method comprising the steps of;

6

7 (a) mounting a camera and a plurality of sensors on
8 a platform capable of movement in the fluid;

9 (b) moving the platform through the fluid while
10 recording visual images on the camera and
11 taking sensor data relating to the attitude and
12 distance of the platform from objects of
13 interest within the fluid;

14 (c) synchronising the visual images to the sensor
15 data to provide corrected visual images
16 relating to a fixed distance and attitude;

17 (d) video mosaicing the images to form an accurate
18 video mosaic as a visual image of the scene
19 surveyed.

20

21 21. A method as claimed in Claim 20 wherein the method
22 includes the step of precalibrating the camera to
23 compensate for distorting artefacts inherent within
24 the camera.

25

26 22. A method as claimed in Claim 20 or 21 wherein the
27 method includes the step of displaying the visual
28 image.

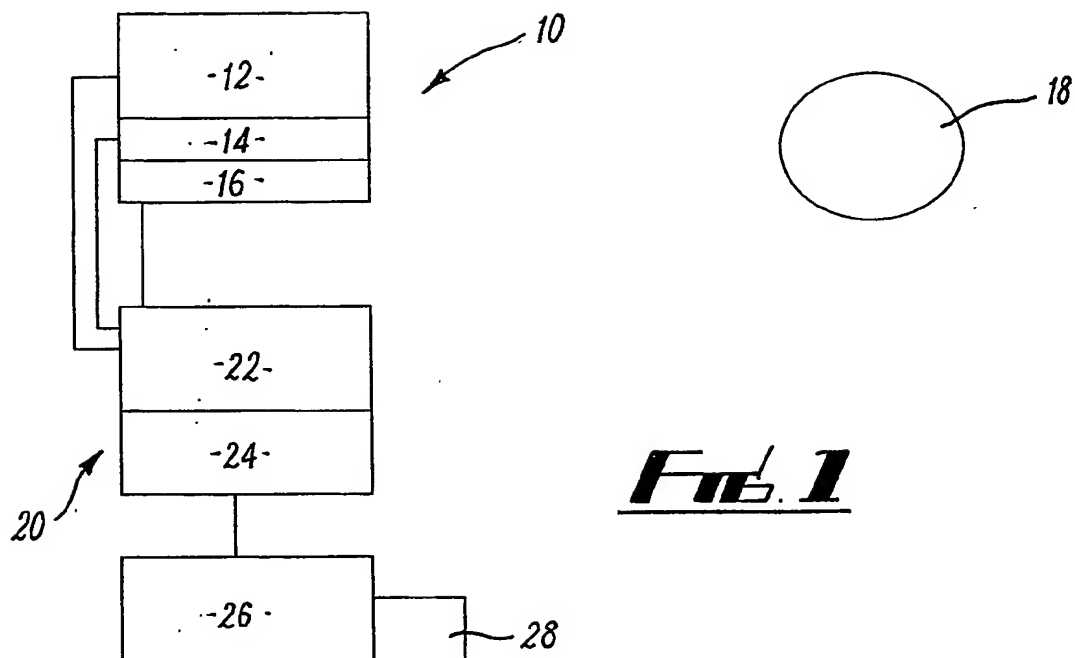
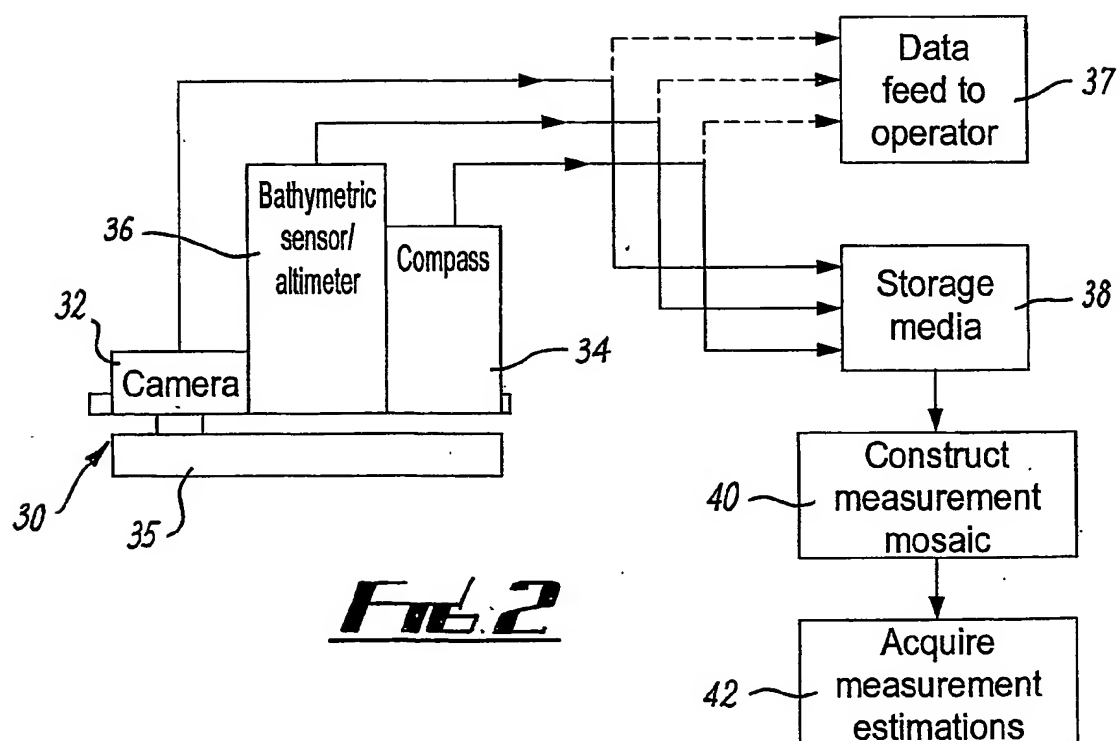
29

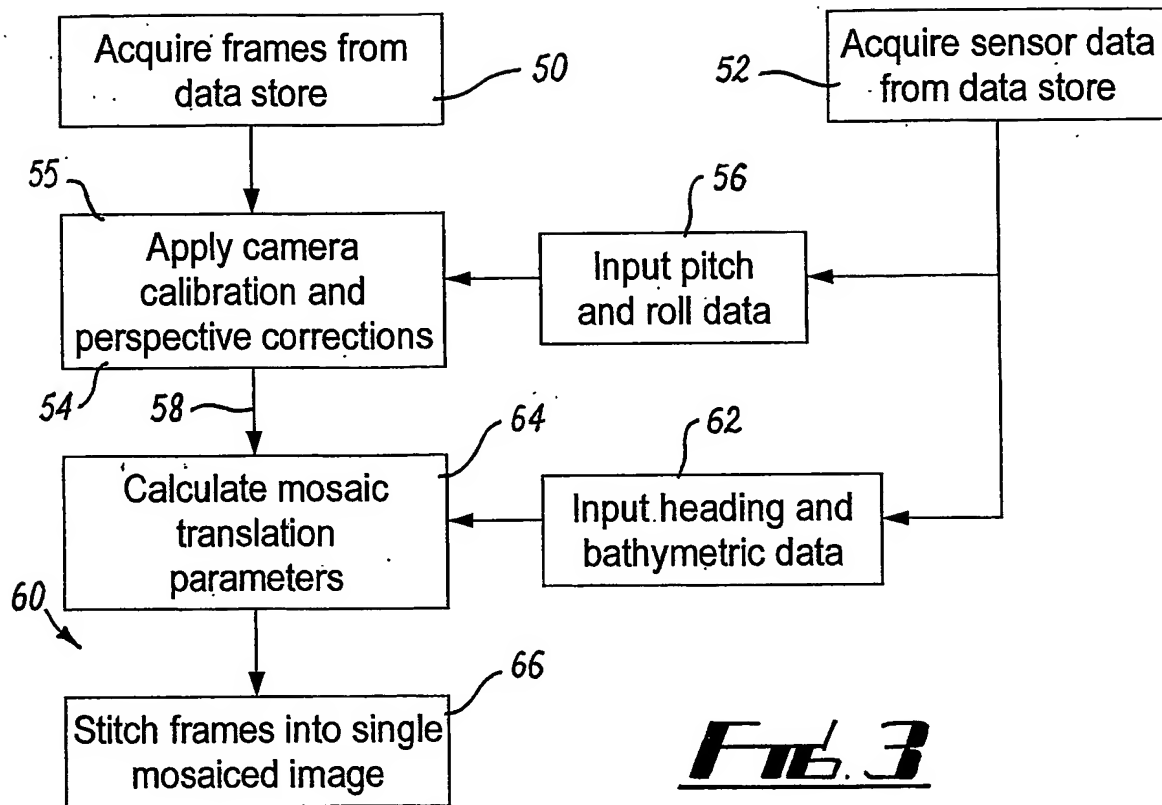
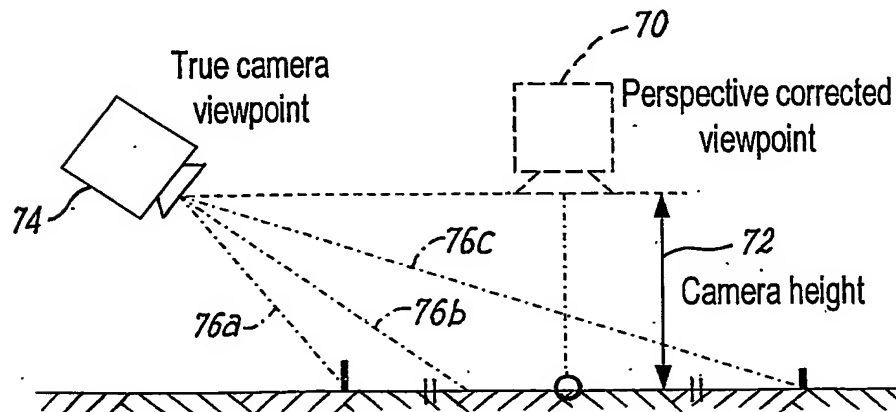
30 23. A method as claimed in any one of Claims 20 to 22
31 wherein the method includes the step of taking a
32 measurement from the visual image.

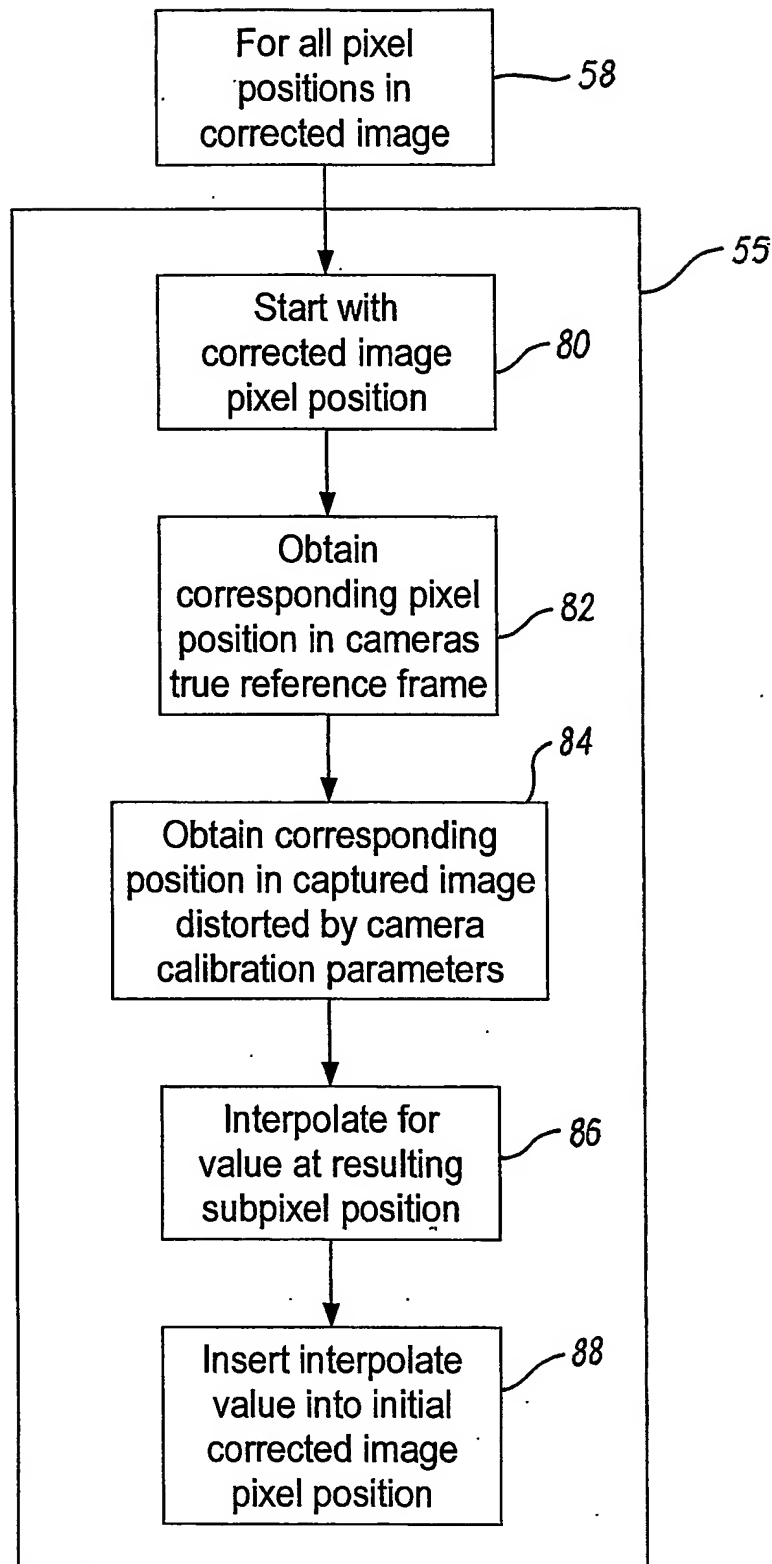
33

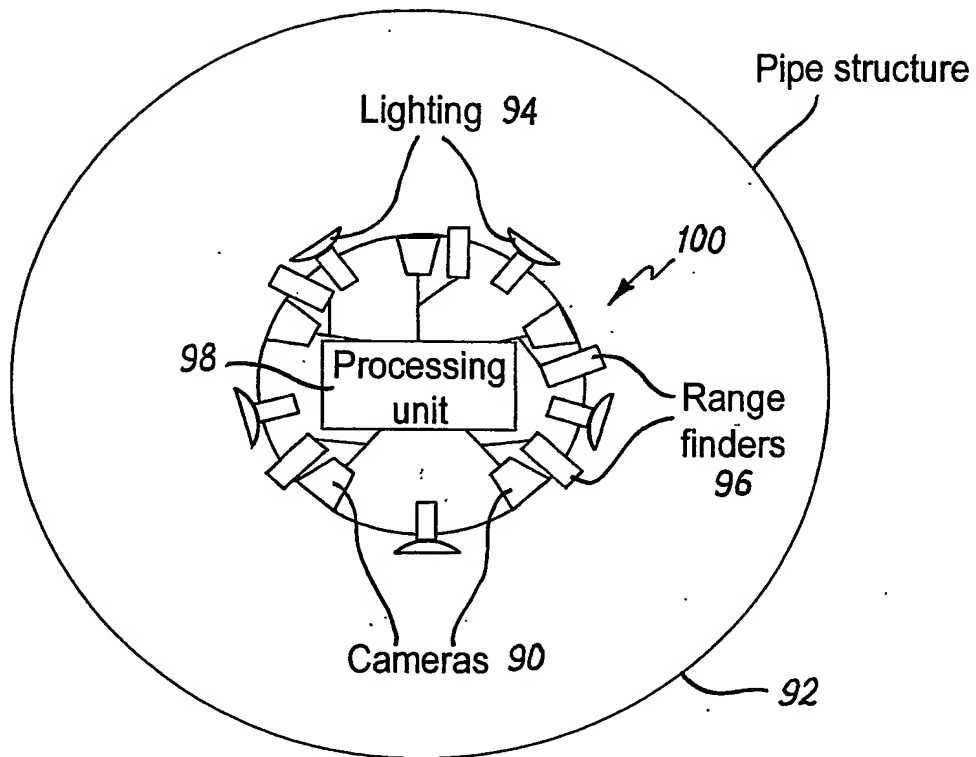
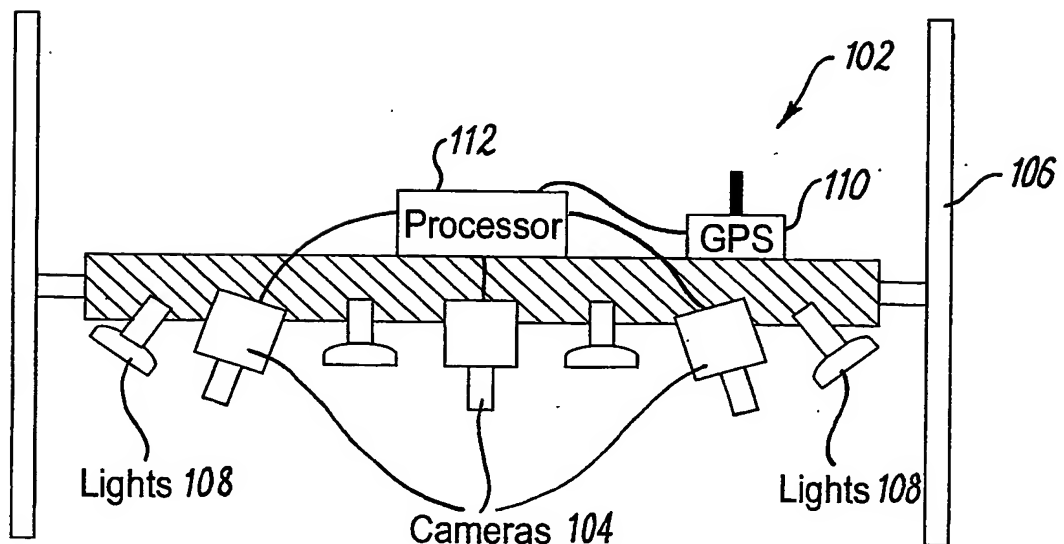
1 24. A method as claimed in any one of Claims 20 to 23
2 wherein the platform is mounted on a remotely
3 operated vehicle (ROV).
4

5 25. A method as claimed in any one of Claims 20 to 24
6 wherein the method includes the step of storing the
7 mosaiced images for viewing later.

**FIG. 1****FIG. 2**

***FIG. 3******FIG. 4***

***FIG. 5***

**Fig. 6****Fig. 7**

INTERNATIONAL SEARCH REPORT

PCT/GB 03/04163

A. CLASSIFICATION OF SUBJECT MATTER IPC 7 G06T7/00 G06T5/50		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 7 G06T G01N G03B		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) EPO-Internal, WPI Data, PAJ, INSPEC		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2001/026684 A1 (FRIDENTAL RON ET AL) 4 October 2001 (2001-10-04) abstract; claims 1-7; figure 2	1, 3, 7-19
A	paragraph '0005! - paragraph '0012!	2, 4-6, 20-25
X	US 2001/038718 A1 (HSU STEPHEN CHARLES ET AL) 8 November 2001 (2001-11-08) abstract; figures 1, 2, 9	1-3, 5-19
A	paragraph '0029! - paragraph '0048!	4, 20-25
X	EP 0 462 905 A (SONY CORP) 27 December 1991 (1991-12-27)	1-3, 5-19
A	abstract; claim 1; figures 1, 2	4, 20-25
-/--		
<input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. <input checked="" type="checkbox"/> Patent family members are listed in annex.		
* Special categories of cited documents : "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 19 February 2004		Date of mailing of the international search report 04/03/2004
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016		Authorized officer Herter, J

INTERNATIONAL SEARCH REPORT

PCT/GB 03/04163

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 5 963 664 A (ANANDAN PADMANABHAN ET AL) 5 October 1999 (1999-10-05) abstract; figure 1 column 4, line 56 - column 5, line 17 -----	1-25
A	US 6 389 179 B1 (KATAYAMA TATSUSHI ET AL) 14 May 2002 (2002-05-14) abstract; figure 33 column 32, line 31 - column 33, line 60 -----	1-25
A	US 6 434 280 B1 (ZOMET ASSAF ET AL) 13 August 2002 (2002-08-13) abstract; figures 2,3A,3B,3C column 2, line 49 - column 3, line 7 -----	1-25
A	GUERMESTEKIN S ET AL: "Image registration and mosaicing using a self-calibrating camera" IMAGE PROCESSING, 1998. ICIP 98. PROCEEDINGS. 1998 INTERNATIONAL CONFERENCE ON CHICAGO, IL, USA 4-7 OCT. 1998, LOS ALAMITOS, CA, USA, IEEE COMPUT. SOC, US, 4 October 1998 (1998-10-04), pages 818-822, XP010308698 ISBN: 0-8186-8821-1 the whole document -----	1-25
A	SALVI J ET AL: "A comparative review of camera calibrating methods with accuracy evaluation" PATTERN RECOGNITION, PERGAMON PRESS INC. ELMSFORD, N.Y, US, vol. 35, no. 7, July 2002 (2002-07), pages 1617-1635, XP004345158 ISSN: 0031-3203 the whole document -----	1-25
A	SAWHNEY H S ET AL: "TRUE MULTI-IMAGE ALIGNMENT AND ITS APPLICATION TO MOSAICING AND LENS DISTORTION CORRECTION" IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, IEEE INC. NEW YORK, US, vol. 21, no. 3, March 1999 (1999-03), pages 235-243, XP000833459 ISSN: 0162-8828 the whole document -----	1-25

INTERNATIONAL SEARCH REPORT

PCT/GB 03/04163

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
US 2001026684	A1	04-10-2001	NONE		
US 2001038718	A1	08-11-2001	US	6512857 B1	28-01-2003
			AU	4561201 A	24-09-2001
			CA	2402929 A1	20-09-2001
			EP	1261940 A2	04-12-2002
			WO	0169535 A2	20-09-2001
EP 0462905	A	27-12-1991	JP	3021556 B2	15-03-2000
			JP	4052635 A	20-02-1992
			CA	2044997 A1	21-12-1991
			DE	69124146 D1	27-02-1997
			DE	69124146 T2	07-05-1997
			EP	0462905 A2	27-12-1991
			KR	184245 B1	15-05-1999
			US	5262867 A	16-11-1993
US 5963664	A	05-10-1999	CA	2225400 A1	09-01-1997
			EP	0842498 A2	20-05-1998
			JP	11509946 T	31-08-1999
			WO	9701135 A2	09-01-1997
US 6389179	B1	14-05-2002	JP	9322040 A	12-12-1997
			JP	9322055 A	12-12-1997
			US	2002126890 A1	12-09-2002
US 6434280	B1	13-08-2002	AU	1048399 A	31-05-1999
			WO	9924936 A1	20-05-1999